Super Sampling of Digital Video

22 February 1999

J. Schuler, D. Scribner, M. Kruer
Naval Research Laboratory, Code 5636
Washington, D.C. 20375

ABSTRACT

Certain imaging applications can exploit features beyond the spatial resolution of a sensor through computational super-resolution techniques. One particular method, super-sampling, involves assimilating multiple frames of digital video into composite video frames of higher average sampling density than the original sensor. Under certain conditions of scene motion and sensor characteristics, the effective resolution is significantly enhanced.

1. Preliminary on Focal Plane Array Image Sampling

An ideal scene function $\Psi(x_1, x_2, t)$ describes with limitless detail the radiant intensity of the world relative to camera coordinates. An ideal lens geometrically projects this function onto the focal plane,

$$s_{geometric}(x_1, x_2, t) \propto \Psi\left(\frac{x_1}{m_1}, \frac{x_2}{m_2}, t\right)$$

where $[m_1, m_2]$ are the magnification factors. In practice, lenses exhibit limited spatial fidelity and the projected optical image is sampled by array of finite sized detectors. The resulting signal is degraded due to the separable effects of lens and pixel to generate a *pseudo-image,*[1] expressed in the space and spatial frequency domains as

$$s_{pseudo}(x_1, x_2, t) = h_{pixel}(x_1, x_2) ** h_{psf}(x_1, x_2) ** s_{geometric}(x_1, x_2, t)$$
$$S_{pseudo}(F_1, F_2, t) = H_{pixel}(F_1, F_2) H_{psf}(F_1, F_2) S_{geometric}(F_1, F_2, t)$$

The spatial extent of the photo-detector pixel with dimensions $[a_1, a_2]$ serves as a "box car" integrator of the geometric image with impulse response

$$h_{pixel}(x_1, x_2) = rect\left(\frac{x_1}{a_1}\right) rect\left(\frac{x_2}{a_2}\right)$$

Form SF298 Citation Data

Report Date *("DD MON YYYY")* 22021999	**Report Type** N/A	**Dates Covered (from... to)** *("DD MON YYYY")*

Title and Subtitle Super Sampling of Digital Video	**Contract or Grant Number**
	Program Element Number
Authors	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Naval Research Laboratory, Code 5636 Washington, D.C. 20375	**Performing Organization Number(s)**
Sponsoring/Monitoring Agency Name(s) and Address(es)	**Monitoring Agency Acronym**
	Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes	
Abstract	
Subject Terms	
Document Classification unclassified	**Classification of SF298** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** unlimited
Number of Pages 6	

The spatial frequency response of such a pixel is the separable function[2]

$$H_{pixel}(F_1, F_2) = a_1 a_2 \left(\frac{\sin(a_1 F_1)}{a_1 F_1} \right) \left(\frac{\sin(a_2 F_2)}{a_2 F_2} \right)$$

The lens point spread function can be characterized in the spatial frequency domain as a monotonically decreasing function.

Assuming a distortion-less optic, the Modulation Transfer Function (MTF) of a unit-sized pixel in one direction of continuous spatial frequency is graphically presented in Figure 1.

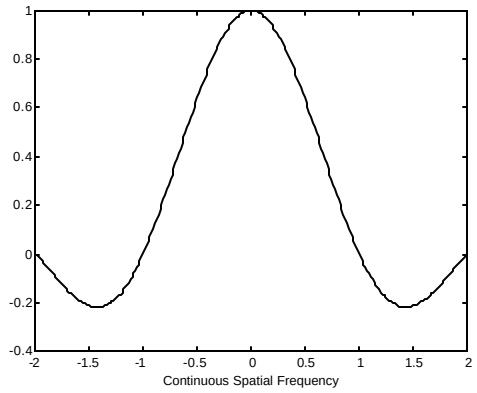

Figure 1 Continuous Space MTF of a Unit Sized Pixel

Integrating and reading out the photo-detector array with detector pitches $[b_1, b_2]$ at a clock rate t_{clock} digitizes the pseudo-image, creating a discrete representation

$$s_{digital}[n_1, n_2, n_3] = s_{pseudo}(n_1 b_1, n_2 b_2, n_3 t_{clock})$$

The Discrete Space Fourier Transform of the sampled image is a periodic extension of the Continuous Space Fourier Transform of the pseudo-image[3]

$$S_{digital}(f_1, f_2, n_3) = \frac{1}{b_1} \frac{1}{b_2} \sum_k \sum_l S_{pseudo}\left(\frac{f_1 - k}{b_1}, \frac{f_2 - l}{b_2}, n_3 t_{clock} \right)$$

where discrete frequency is related to continuous frequency by

$$f_k = \frac{F_k}{F_{sample}} = b_k F_k$$

The MTF of the same unit sized pixel sampled in a 100% fill factor lattice is graphically presented in Figure 2 for one direction of discreet spatial frequency. Note that the low spatial sampling rate of a FPA introduces significant alias distortion of all but the lowest spatial frequencies.

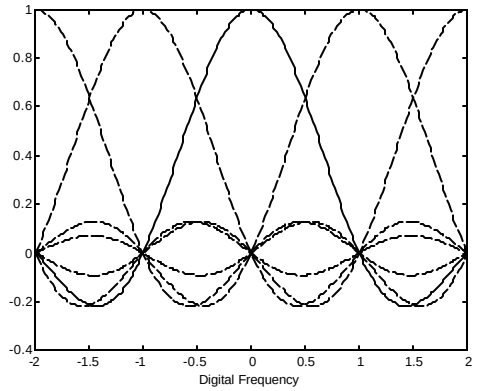

Figure 2 Discrete Space MTF of a 100% Fill Factor FPA

Conventionally, spatial resolution is significantly improved by introducing a blurring *Nyquist* optic to pre-filter all spatial frequencies above the sampling fold-over frequencies at $\pm\frac{1}{2}$.

Clever alternatives to such a blurring optic is to dither or micro-scan the FPA imaging system to synthetically increase the sampling rate. The MTF that results from a 6-fold reduction in detector pitches $[b_1, b_2]$ is graphically presented in Figure 3 for one direction of discreet spatial frequency.

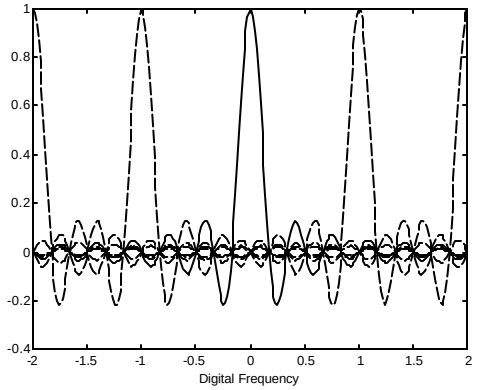

Figure 3 Discrete Space MTF of a 100% Fill Factor FPA subject to 6-fold dithering

Note that although the pixel blur is enhanced over the domain $[-½,+½]$, alias distortion is significantly reduced, permitting restoration by an image-sharpening filter that was inapplicable to the low-resolution image.

2. Motion Estimation

All forms of super-sampling require re-assembly of a temporal cube of digital video into a single image with a higher spatial sampling density[4]. The success of such processing hinges on accurate knowledge of the FPA motion relative to the object scene. Initial approaches that incorporated a mechanically controlled dither held hope that the inter-image motion would be a known uniform shift determined by a commanded position change[5]. Disappointingly, empirical results show that a known-shift model is insufficient to characterize the FPA motion when the sensor platform has additional vibration or translational motion[6]. The next iteration is to estimate the true shift by any member of a well-developed class of image displacement estimation algorithms[7]. Even still, this is insufficient to characterize the motion for most image sequences acquired from moving platforms[8]. As previously reported, we generated very promising results by applying numerous image displacement estimators across the image and fitting

the resulting motion field to an affine displacement model characteristic of classical perspective transformations. Current research is exploring wavelet based parametric models[9] to account for higher order terms of the basic perspective transformation to account for simple geometric distortions of the sensor optics.

3. Results

We extend our analysis from 128x128 MWIR data of the NRL Threat Warning Program[10] to 640x480 MWIR data of the previously reported SASSY program[11]. We were provided several short (30-50 frame) sequences from an airborne ingress into a collection of targets. For compliance with RS-170 video, a 300x480 window of imagery with a 2:1 interlace was digitally recorded. Hence, every FPA integration window alternated between even and odd subsets of the 480 available rows to generate a new 300x240 image.

Because of the very limited temporal data, we could not compute a fully meaningful affine optical flow model as previously demonstrated. Instead we utilized a strictly translational displacement model with limited success, and implemented a 5-fold increase in the spatial sampling rate of the video sequence.

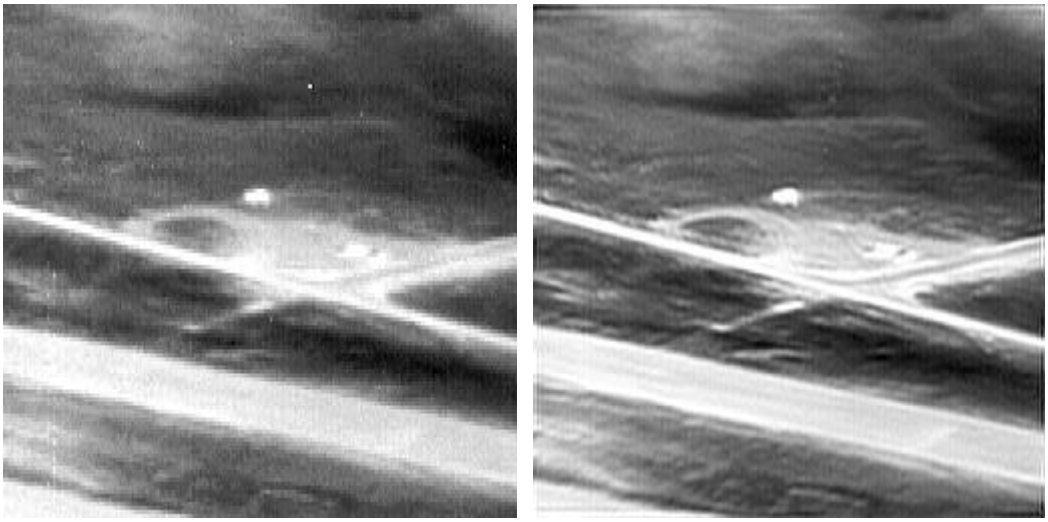

Figure 4 128x128 window of Data (Left) and subsequent 5-fold Super Sampling (Right)

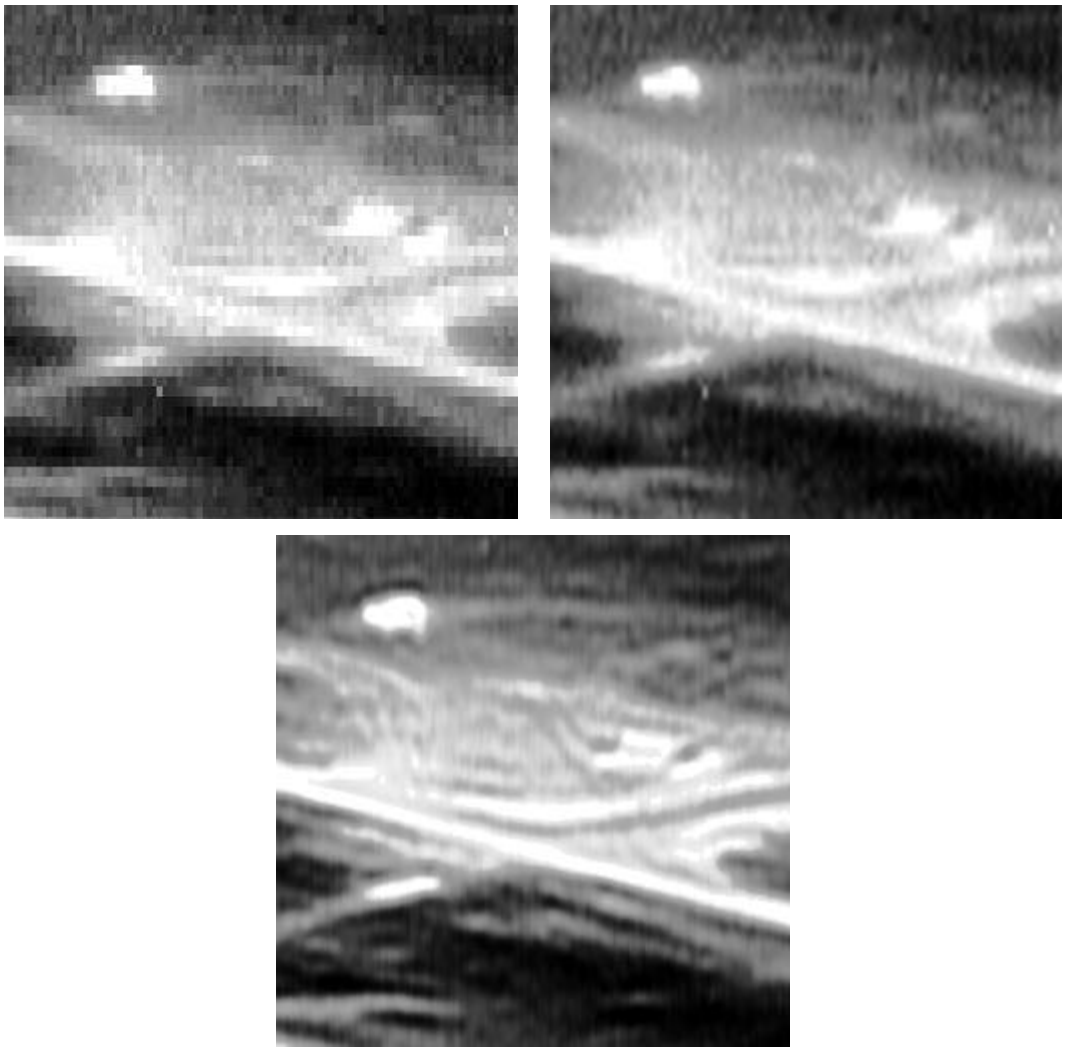

Figure 5 Original Data (Upper Left), Interpolated Imagery (Upper Right) Super Sampled Imagery (Bottom)

4. Conclusion

Resolution enhancement through temporal super sampling has been demonstrated on a number of recorded digital video sequences. Most significantly, none of the video sequences involved controlled micro-dither scanning, and no consideration was given for the subsequent resolution enhancement analysis performed years after data acquisition. This underscores the fundamental robustness of the super sampling approach, and suggests an immediate deployment in a prototype real-time system without further need for the development of micro-dithering hardware.

---

[1] D.A. Scribner, M.S. Longmire, M.R. Kruer, "Analytic modeling of Staring Infrared Systems with Multidimensional Matched Fitlers", SPIE Vol. 890 Infrared Systems and Components II (1988)

[2] G.R. Cooper, C.D. McGillem, Modern Communications and Spread Spectrum, McGraw-Hill Inc.(1986)

[3] J.G. Proakis, D.G. Manolakis, Digital Signal Processing Principles, Algorithms, and Applications, Macmillan Publishing Company (1992)

[4] J.M. Schuler, D.A. Scribner, M.R. Kruer, "Super Resolution Imagery from Multi-Frame Sequences with Random Motion" Proceedings on the 1998 Meeting of the IRIS Specialty Group on Passive Sensors Vol. 1 (1998)

[5] W.F. O'Neal "Experimental Performance of a Dither-Scanned InSb Array" " Proceedings on the 1993 Meeting of the IRIS Specialty Group on Passive Sensors (1993)

[6] S. Cain, E. Armstrong, B. Yasuda "Joint Estimation of Image, Shifts, and NonUniformities from IR Images" " Proceedings on the 1997 Meeting of the IRIS Specialty Group on Passive Sensors Vol. 1 (1997)

[7] A. Schaum, M. McHugh, "Analytic Methods of Image Registration: Displacement Estimation and Resampling", NRL Report 9298 U.S. Naval Research Laboratory, Washington, DC (1992)

[8] Q. Zheng, R. Chellappa "A Computational Vision Approach to Image Registration" Tech. Rep. CAR-TR-583 Center for Automation Research, University of Maryland (1991)

[9] J. Magerey, N. Kingsbury "Motion Estimation Using a Complex-Valued Wavelet Transform" IEEE Transactions on Signal Processing Vol. 46, No4 (April 1998)

[10] K. Sarkedy et. al. "Performance Evaluation of a Staring Two Color Missile Approach Warning Sensor" Proceedings of the 1995 IRIS Specialty Group on Passive Sensors (1995)

[11] S.B. Campana, B. Haugh, J.Toner, L. McMichael "Results of Airborne Tests of a Medium Wavelength Infrared Imager" Proceedings of the 1996 Meeting of the IRIS Specialty Group on Passive Sensors (1996)